SYSTEM AND METHOD FOR DYNAMICALLY

ADAPTING A BANNER ADVERTISEMENT

TO THE CONTENT OF A WEB PAGE

5

FIELD OF THE INVENTION

The present invention relates to the field of data processing. Particularly, this invention relates to a software system and associated method for use in e-commerce advertising with a search engine that searches data maintained in systems that are linked together over an associated network such as the Internet. More specifically, this invention pertains to a computer software product for dynamically adapting a banner advertisement to the categorization, surrounding page content, and changes of the advertiser's repository.

BACKGROUND OF THE INVENTION

The World Wide Web (WWW) is comprised of an expansive network of interconnected computers upon which businesses, governments, groups, and individuals throughout the world maintain inter-linked computer files known as web pages. Users navigate these pages by means of computer software programs

20 commonly known as Internet browsers. Due to the vast number of WWW sites, many web pages have a redundancy of information or share a strong likeness in either function or title. The vastness of the unstructured WWW causes users to rely primarily on Internet search engines to retrieve information or to locate businesses. These search

engines use various means to determine the relevance of a user-defined search to the information retrieved.

The authors of web pages provide information known as metadata, within the body of the hypertext markup language (HTML) document that defines the web pages. A computer software product known as a web crawler, systematically accesses web pages by sequentially following hypertext links from page to page. The crawler indexes the pages for use by the search engines using information about a web page as provided by its address or Universal Resource Locator (URL), metadata, and other criteria found within the page. The crawler is run periodically to update previously stored data and to append information about newly created web pages. The information compiled by the crawler is stored in a metadata repository or database. The search engines search this repository to identify matches for the user-defined search rather than attempt to find matches in real time.

A typical search engine has an interface with a search window where the user enters an alphanumeric search expression or keywords. The search engine sifts through available web sites for the user's search terms, and returns the search of results in the form of HTML pages. Each search result includes a list of individual entries that have been identified by the search engine as satisfying the user's search expression. Each entry or "hit" may include a hyperlink that points to a Uniform Resource Locator (URL) location or web page.

In addition to the hyperlink, certain search result pages include a short summary or abstract that describes the content of the URL location. Typically, search engines generate this abstract from the file at the URL, and provide acceptable results for URLs that point to HTML format documents. For URLs that point to HTML documents or web

5   pages, a typical abstract includes a combination of values selected from HTML tags. These values may include a text from the web page's "title" tag, from what are referred to as "annotations" or "meta tag values" such as "description", "keywords", etc., from "heading" tag values (e.g., HI, H2 tags), or from some combination of the content of these tags.

More specifically, the popularity of portal sites, that act as gateways to very specialized information sources, has grown concurrently with the WWW, both in complexity and volume of data. The term "portal" is generally synonymous with gateway, and is typically used to refer to a WWW site which is intended to be a major

5   starting site or as an anchor site for web users. Current leading general-purpose portal sites include: Yahoo!®, Excite®, Netscape®, Lycos®, Cnet®, and MSN The Microsoft Network®. However, while such portal sites attempt to serve as gateways to a wide variety of general-purpose information, specialized portals have also been gaining popularity in recent years.

20

The portal database is a vast repository of pre-collected, indexed, and summarized information, typically gathered from the WWW using automated crawling tools. When a user enters a query, the portal's search engine attempts to match the keywords

ARC9-2000-0100-US1   3

specified by the user with summarized metadata that have been previously extracted from the documents stored in the repository, and then returns an ordered list of matches relevant to the user's query.

5 Typically, the search engine will return a result set for a search query including a URL and a text based abstract of the original resource. Sometimes, users are able to control the length of the abstract. For instance, the HotBot® site at URL: http://www.hotbot.com, provides the choice of having only a list of URLs displayed as the search result, the URL with a brief abstract, or a comprehensive abstract.

10 Currently, many web pages contain advertisements that assume various forms such as banner ads (or advertisements) across the top or bottom of the page. Such ads may include scrolled information containing images that change with time. Disadvantageously, from an advertiser's perspective, web users have a tendency to

15 mentally "tune-out" such advertisements as they read or interact with the information displayed on the main work area of a page. Furthermore, by utilizing a portion of the valuable "real estate" on a web page for advertisement, the remaining available work area on the page is reduced from its maximum full-screen capabilities.

20 Banner ads can have text, still or moving graphics, or multimedia messages, and typically serve as hypertext links, such that the user is linked to other specified pages if the user clicks on the banner ads. Banner ads can be categorized as corporate image ads, and information ads. The main purpose of corporate image ads is to enhance the

ARC9-2000-0100-US1   4

visibility and public image of a business enterprise, and to reflect its presence, participation and involvement in a particular domain.

5   The information banner ads highlight a specific product, service, or content, and provide a URL to corresponding content information pages. The context placement of these banner ads is critical in that it needs to match the interests of potential customers. Currently, advertisers are able to select the surrounding content of the banner ads based primarily on the content categories.

10   For example, a developer portal wishes to advertise on a search service provider such as Yahoo!® in order to gain more traffic. Search service providers offer a variety of categories where to place product or content ads. As an illustration, "software development", "Java®", "XML", etc. might constitute reasonable categories for an ad placement for the developer portal. To place the same ad within a "Home & Garden"

15   category would be a misplacement, since the percentage of potential customers who are simultaneously seeking home and garden products and a software development product is not high.

       Such misplacement is a common occurrence due in part to the static nature of the

20   banner ad. Reference is made to the following publications:

       U.S. Patent No. 6,009,410, titled "Method and System for Presenting Customized Advertising to a User on the World Wide Web";

       U.S. Patent No. 6,014,502, titled "Electronic Mail System with Advertising"; and

ARC9-2000-0100-US1   5

U.S. Patent No. 5,937,392, titled "Banner Advertising Display System and Method with Frequency of Advertisement Control".

The static content of the banner ads might be acceptable as corporate image ads

5   where the goal of the ad is to spread the visibility of a company. However, as products and services of a company continue to change, it would be advantageous to have the banner ads automatically reflect these changes.

As an example, for a data store carrying a variety of products and services, it would

10   be desirable to have the newer or top rated products and services within specific categories automatically updated and displayed in the banner ads. Currently, the most viable approach is for the advertiser to manually update the banner ads to reflect the desired products and services. However, such a "static approach" presents several disadvantages, among which are the following:

15   a) the selection might become obsolete after a short period of time; and

           b) the maintenance effort to administer and manage the banner ads will be too high to support over an extended period of time. In particular, the problem of maintaining the banner ads content up to date becomes increasingly difficult for companies that provide a variety of different products and multimedia information within a repository that

20   continuously changes over a short time interval.

There is currently no adequate mechanism by which the content of the banner ads is automatically updated based on content changes of the advertiser's repository. Such

ARC9-2000-0100-US1   6

adaptive process would provide up to date content and information. The need for such an adaptive mechanism and corresponding process has heretofore remained unsatisfied.

5   ## SUMMARY OF THE INVENTION

The adaptive advertising system and associated method of the present invention satisfy this need. In accordance with one embodiment, the adaptive ad system dynamically adapts the content of a banner ad to the categorization, surrounding page content, and changes of the advertiser's repository of products and services.

10

In addition, the adaptive advertising system provides appropriate information resources based on the user's needs. As an example, IBM®, as an advertiser wishes to place a banner advertisement for IBM® developerWorks under the "Java" category of Yahoo!®. The adaptive advertising system recognizes the context of this advertisement, namely "Java," and provides the top content URLs within the IBM® developerWorks Java zone. If the same advertisement were to be placed within the "Java/EJB" Yahoo!® category, the adaptive advertising system will automatically recognize the more specific (or specialized) context of the advertisement and will provide the top URLs within the IBM® developerWorks Java zone, with a focus on EJB.

20

As a result, the adaptive advertising system provides the capability to serve advertisements with adaptive contents. This level of adaptivity ensures that the content of the banner advertisement reflects the current content of the web page where it is

ARC9-2000-0100-US1   7

embedded, with a high degree of confidence. Advertisers using advertisements with adaptive content are relieved from the tedious and time and resource consuming task of having to repeatedly create new advertisements that are specifically designed for different page contents. The adaptive advertising system will automatically adapt the advertisement to the continuously changing page content.

Even within a category, for instance the "Java" category in the example above, the overall content of the page might not precisely reflect this category. This situation occurs when a content provider does not have control over the content within a category. A popular content provider of this genre is "Deja.com" with its Usenet discussions. Although a category like "comp.programming" promises programming content, the discussion content could be about an unrelated topic, such as IBM's AS/400, which would be more appropriately categorized as a propriety hardware. This illustration shows that in the situation where a content provider looses control over the page content, the adaptive advertising system analyzes the page content and adapts the actual content of the banner advertisement to the page content.

Based on the page content, the adaptive advertising system determines whether or not to display the banner advertisement. If the content is inappropriate, the adaptive advertising system might decide not to display the banner advertisement to avoid an undesirable association between the banner advertisement and the page content. As an illustration, consider an IBM advertisement within the "comp.programming" category being displayed nest to an article with an obscene content. IBM's corporate image might

ARC9-2000-0100-US1 8

not be well served with such an undesirable association. The adaptive advertising system identifies this scenario, and disables itself, i.e., prevents the display of the banner ad, to avoid such a negative image association.

5   Therefore, the adaptive advertising system either displays or suppresses the banner ad based on the surrounding page content. This involves taking any one or more of the following steps:

a) Fine tuning the advertisement by showing the advertisement in the proper specialized category. For example, for a category "Java" with a surrounding page content "EJB", the advertisement content will incorporate Java with focus on EJB.

b) Replacing the category content. For example, for a category "Java" with a surrounding page content "AS/400", the advertisement content will focus only on AS/400.

c) The adaptive advertising system disables the advertisement until such time as the page content changes.

The foregoing and other features and advantages of the present invention are realized by an adaptive advertising system that can be used in the context of an Internet environment. Transparently to the user, the system continuously operates in the background to adapt banner advertisements based on the page content, surrounding

content, and specific categorization or keywords provided by a domain specific repository.

The system is generally comprised of a banner display module, a keyword analyzer, an ad proxy router, an ad server, a banner advertising manager, an ad search engine, an indexer, an ad repository, an ad index repository, an advertiser site repository, and optionally a domain specific repository.

The keyword analyzer analyzes the page content, and the banner display module determines the desirability of associating the advertisement with the page. If the banner display module determines that such an association does not adversely impact the advertiser's image, the banner display module selectively displays the advertisement. Otherwise, the banner display module suppresses the advertisement.

In one embodiment, the advertisement includes a static portion such as the advertiser's logo, and a dynamic portion. The dynamic portion can be any one or more of: multimedia files; advertisements, executable codes, or hypertext links.

The banner display module sends a data stream containing the following information to the proxy router: the selected category; the keyword from the page; and the address of the ad server. In turn, the ad proxy router sends the following information to the ad server: the session information; the selected category; and the keywords from the page.

The indexer indexes the content of the advertiser's site, and stores the generated

hyperlinks in the ad index repository. The ad repository stores the following: various

advertisements from the advertiser; multimedia files; and executable codes or

applications.

5

BRIEF DESCRIPTION OF THE DRAWINGS

The various features of the present invention and the manner of attaining them will

be described in greater detail with reference to the following description, claims, and

drawings, wherein reference numerals are reused, where appropriate, to indicate a

10   correspondence between the referenced items, and wherein:

FIG. 1 is a schematic illustration of an exemplary operating environment in which an

adaptive advertising system of the present invention can be used;

FIG. 2 is a more detailed block diagram of the adaptive advertising system of FIG. 1

15   shown implemented in part on the user's side;

FIG. 3 is a more detailed block diagram of the adaptive advertising system of FIG. 1

shown implemented in part on a server;

20

FIG. 4 is a flow chart that depicts the operation of the adaptive advertising system of

FIGS.1 - 3.

DETAILED DESCRIPTION OF THE INVENTION

The following definitions and explanations provide background information pertaining to the technical field of the present invention, and are intended to facilitate the understanding of the present invention without limiting its scope:

5

Banner advertisement (or ad): A message, usually but not necessarily displayed for a fee, and associated with products and/or services offered by an advertiser.

Crawler: A program that automatically explores the World Wide Web by retrieving a document and recursively retrieving some or all the documents that are linked to it.

Dictionary: A database of context-related terms. A domain specific dictionary includes domain specific repositories such as a dictionary, a thesaurus, and other similar data stores.

15

HTML (Hypertext Markup Language): A standard language for attaching presentation and linking attributes to informational content within documents. During a document authoring stage, HTML "tags" are embedded within the informational content of the document. When the web document (or "HTML document") is subsequently transmitted by a web server to a web browser, the tags are interpreted by the browser and used to parse and display the document. In addition to specifying how the web browser is to display the document, HTML tags can be used to create hyperlinks to other web documents.

ARC9-2000-0100-US1   12

Internet: A collection of interconnected public and private computer networks that are linked together with routers by a set of standards protocols to form a global, distributed network.

5

Search engine: A remotely accessible World Wide Web tool that allows users to conduct keyword searches for information on the Internet.

Server: A software program or a computer that responds to requests from a web browser by returning ("serving") web documents.

10

URL (Uniform Resource Locator): A unique address that fully specifies the location of a content object on the Internet. The general format of a URL is protocol://server-address/path/filename.

15

Web browser: A software program that allows users to request and read hypertext documents. The browser gives some means of viewing the contents of web documents and of navigating from one document to another.

20 Web document or page: A collection of data available on the World Wide Web and identified by a URL. In the simplest, most common case, a web page is a file written in HTML and stored on a web server. It is possible for the server to generate pages dynamically in response to a request from the user. A web page can be in any format

ARC9-2000-0100-US1   13

that the browser or a helper application can display. The format is transmitted as part of the headers of the response as a MIME type, e.g. "text/html", "image/gif". An HTML web page will typically refer to other web pages and Internet resources by including hypertext links.

5

Web Site: A database or other collection of inter-linked hypertext documents ("web documents" or "web pages") and associated data entities, which is accessible via a computer network, and which forms part of a larger, distributed informational system such as the WWW. In general, a web site corresponds to a particular Internet domain name, and includes the content of a particular organization. Other types of web sites may include, for example, a hypertext database of a corporate "intranet" (i.e., an internal network which uses standard Internet protocols), or a site of a hypertext system that uses document retrieval protocols other than those of the WWW.

World Wide Web (WWW): An Internet client - server hypertext distributed information retrieval system.

FIG. 1 portrays the overall environment in which an adaptive advertising system 10 according to the present invention may be used. The system 10 includes a software or computer program product which is typically embedded within, or installed, at least in part, on a host server 15. Alternatively, the system 10 can be saved on a suitable storage medium such as a diskette, a CD, a hard drive, or like devices. While the system 10 will be described in connection with the WWW, the system 10 can be used

ARC9-2000-0100-US1   14

with a stand-alone database of documents that may have been derived from the WWW and/or other sources.

The cloud-like communication network 20 is comprised of communication lines and switches connecting servers such as servers 25, 27, to gateways such as gateway 30. The servers 25, 27 and the gateway 30 provide the communication access to the WWW Internet. Users, such as remote Internet users are represented by a variety of computers such as computers 35, 37, 39, and can query the host server 15 for the desired information.

The host server 15 is connected to the network 20 via a communications link such as a telephone, cable, or satellite link. The servers 25, 27 can be connected via high speed Internet network lines 44, 46 to other computers and gateways. The servers 25, 27 provide access to stored information such as hypertext or web documents indicated generally at 50, 55, and 60. The hypertext documents 50, 55, 60 most likely include embedded hypertext link to other locally stored pages, and hypertext links 70, 72, 74, 76 to other webs sites or documents 55, 60 that are stored by various web servers such as the server 27.

FIGS. 2 and 3 illustrate a high level architecture showing the adaptive advertising system 10 used in the context of an Internet environment. The system 10, transparently to the user, continuously or periodically operates in the background to adapt banner advertisements based on the page content, surrounding content, and specific

ARC9-2000-0100-US1   15

categorization or keywords provided by one or more domain specific repositories such as a dictionary, a thesaurus, and so forth.

The system 10 is generally comprised of a banner display module 200, a keyword analyzer 210, an ad proxy router 212, an ad server 214, a banner advertising manager 220, an ad search engine 230, an indexer 252, an ad repository 240, an ad index repository 242, an advertiser site repository 244, and optionally a domain specific dictionary / repository 250.

In operation, and with further reference to FIG. 4, the web page 150 is rendered and displayed to the user at step 405. While the user is browsing the web page 150, the keyword analyzer 210 analyzes the page content at step 410. At decision step 415, the banner display module 200 of the adaptive advertising system 10 determines the desirability of associating this page 150 with the banner advertisement 160 assigned by the server 15 for display in conjunction with the page 150. Although the adaptive advertising system 10 will be described in terms of a single banner advertisement 160, it should be clear that the system 10 is applicable to two or more simultaneous advertisements as well.

If at decision step 415 the adaptive advertising system 10 determines that the banner advertisement 160 will be misplaced if displayed in the current page 150, it will not display the banner ad 160 on the page 160 (step 420), but will await the arrival of the next page at step 425.

ARC9-2000-0100-US1   16

Otherwise, if the adaptive advertising system 10 determines at decision step 415 that the banner advertisement 160 can be displayed without disadvantageously affecting the advertiser's image, the banner display module 200 proceeds to display a first portion, i.e., a static portion 270, of the banner advertisement 160 (step 430). In one embodiment, the static portion 270 can be the advertiser's logo, which is common to most or all of the advertisements stored in the advertiser's ad repository 240.

Simultaneously, the banner display module 200 analyzes the page content using the keyword analyzer 210 and the domain specific dictionary repository 250 at step 440, to select the most appropriate advertisement or advertisements 160 from the ad repository 240 and/or the most appropriate links (i.e., URLs, pointers, hyperlinks, or addresses) from the ad index repository 242.

To this end, the banner display module 200 sends a data stream containing the following information to the web server 15:

- categories and keywords and other relevant information from the keyword analyzer 210; and

- URL or address of the advertiser's site or ad server 214.

In turn, the web server 15 forwards the data stream to the ad proxy router 212 using, for example the single object access protocol (SOAP). The ad proxy router 212 can be an integral part of the web server 15 or, alternatively, it can be a separate component. The ad proxy router 212 uses the advertiser site's URL contained in the data stream,

ARC9-2000-0100-US1   17

removes irrelevant information from the data stream, and routes the remaining information in the data stream to the ad server 214.

The ad server 214 processes the data stream, and forwards it to the banner advertising manager 220. Using this information, the banner advertising manager 220 automatically constructs a query and submits it, at step 445, to the ad search engine 230, for retrieving the most appropriate hits at step 450 (FIG. 3). These hits can be comprised of any one or more of the following:

- URLs pointing to specific pages (or text documents) that are stored in the ad site content repository 244;

- various static ad portions 270, such as logos, that are stored in the ad repository 240; and/or

- various dynamic ad portions 280, such as multimedia content and executable codes, that are stored in the ad repository 240.

To this end, the site content repository 244 contains all the documents related to the products and services offered by the advertiser. These documents are indexed by an indexer 252, and the generated indices are stored in the ad index repository 242.

Upon receiving the query from the banner ad manager 220 (FIG. 2 or 3), the ad search engine 230 searches the ad repository 240, and retrieves the related static portions, and subsequently (or concurrently) it also retrieves the dynamic portions 280,

as well as any related executable code and multimedia information, and transmits the same as a return data stream back to the banner ad manager 220.

Similarly, upon, or shortly after, receiving the query from the banner ad manager 220, the ad search engine 230 searches the ad index repository 242, and retrieves the related URLs and additional property values such as the abstract and other content information associated with the URLs and transmits the return data stream back to the banner ad manager 220. It should be noted that the return data streams from the ad repository 240 and the ad index repository 242 can be timed so that they appear in a desired order in the banner ad 160.

In turn, the banner ad manager 220 forwards the return data streams to the web server 15 and therefrom to the banner displayer 200. In order for the return data streams to be forwarded to the appropriate web server 15 and banner display 200, particularly when the adaptive advertising system 10 is handling requests from several users, the web server 15 is equipped with a session logger that integrates the session information into the data stream to the ad proxy router 212 and the ad web server 214. The session information is retained as part of the return data streams and used by the web server to route these return data streams to the appropriate user site or banner displayer 200. In addition, in order for the banner ad manager 220 (or in certain applications for the ad web server 214) to route the return data streams to the proper server, the session logger integrates the address or URL of the web server 15 (or in

certain application to the ad proxy router 212) in the data streams being forwarded to the ad web server 214.

Upon receipt of the return data streams, the banner displayer 200 displays the banner ad 160 on the page 150 as explained herein (step 455).

Having described the main components of the adaptive advertising system 10 and the environment in which it operates, these components will now be individually described in more detail.

The banner display module 200 displays an up-to-date targeted banner ad 160 on the current page 150. The banner ad 160 shows the top 5 or more hits related to the content of the page 150, and any related multimedia file, within the dynamic portion 280 of the banner ad 160, which includes a "hit list" 285. In a preferred embodiment, the dynamic portion 280 is implemented as an applet in Java®, in combination with scripting languages such as JavaScript® within the context of a web browser such as E Microsoft Internet Explorer® or Netscape Navigator®.

The banner display module 200 analyzes the category content information as input parameters, and outputs a banner ad 160 therefrom. In most likelihood, the marketing staff of the advertiser are familiar with the categories under which the advertiser's products and/or services will be categorized, and it is possible to pass these categories as parameters to the applet.

ARC9-2000-0100-US1   20

. As an example, IBM® developerWorks wishes to advertise its Java related content web site on a web portal. The banner ad 150 will be placed in the Java section of the portal, or within a similar classification (e.g. Programming Languages).

5   According to another embodiment, the banner display module 200 automatically attempts to determine the category of the current page 150, using known or available clustering and automatic classification tools, such as IBM® Intelligent Miner®.

To further refine the categorization of the current page 150, the keyword analyzer 10   210 can be used to watch for certain predetermined or "hot" keywords 287. For example, if EJB (Enterprise Java Beans) is the topic frequently mentioned in the page 150, the selected categories are Java and EJB, and the hit list 285 includes links related to "Java EJB". The refined categories are then sent to the web server 15 using network communication, and in return, an XML formatted document embedded with links and 15   descriptions are sent to the web server 15. The links are then displayed in the dynamic portion 280 of the banner ad 160. An image map based on HTML markup can also be used to show the hits.

The keyword analyzer 210 can be located either on the user side (FIG. 2) or on the 20   server side (FIG 3). It is an algorithm that classifies a web page systematically and refines the categories under an already-given classification. To this end, the key analyzer 210 performs three tasks: (1) it filters out "noise" words; (2) it determines if the words detected as keywords in the current page 150 are related to the current category;

ARC9-2000-0100-US1   21

and (3) refines the final category to be sent to the web server 15. Each of these tasks will now be detailed.

As used herein, the current category is the category currently associated with the page 160. Typically, the current category can be manually defined by the person (e.g. a webmaster), who integrates the adaptive ad 150 into the page content 160. This can be done by passing parameters to the applet as described above.

The keyword analyzer 210 filters out "noise" words by removing words that do not constitute a substantive part of the web page 150, prior to performing the analysis of the web page 150. These "noise" words include for example non indexable words such as: these, are, is, the, to, in, be, there, etc. The specific filtering rule may vary depending on the language used in the document, but the general filtering concept remains the same.

The second task of the keyword analyzer 210 is to determine if the words detected as keywords in the current page 150 are related to the current category. To this end, once the "noise" words are filtered out, a clustering tool such as IBM® Intelligent Miner® can be used to determine whether the current page 150 matches the current category.

The keyword analyzer 210 can use the domain specific dictionary repository 250, which is related to the current category to determine the number of words in the web page 150 that match those in the dictionary repository 250. Based on a predetermined threshold level, the keyword analyzer 210 determines whether or not the page 150

ARC9-2000-0100-US1   22

matches a selected category. The selected category can be manually chosen by passing parameters to the adaptive ad applet, or automatically by the banner display module 200.

5   If the result of this analysis is below the threshold level, the keyword analyzer 210 issues a temporary inactivation command to the banner display module 200 to prevent display of the banner ad 160. As discussed earlier, this functionality can be useful to prevent the inappropriate ad placement.

10   One way to determine the threshold level is to calculate the ratio of the occurrence of matched words to the total number of words in the dictionary repository 250. For example, the programmer or designer of a particular adaptive ad implementation sets the threshold level (as a parameter) to 0.2. If the calculated ratio exceeds this threshold level, the keyword analyzer 210 presumes that the selected category is correct.

15

As an illustration, a page is classified under the Java® category, and a Java® dictionary contains 100 words, such as "Java, EJB, swing, jfc, jini, rmi, ...," the word "Java" appears 20 times, and the word swing appears 10 times. The ratio of the number of occurrence of matched words "Java" and "swing" to the total number of words is 30 to 20   100, which is greater than the threshold level 0.2. As a result, the selected category of the page 150 is considered accurate.

Once the keyword analyzer 210 determines the accuracy of the selected category

ARC9-2000-0100-US1   23

for the current page 150, it start the third task of refining the selected category to be sent as a refined category to the web server 15. The keyword analyzer 210 selects the most likely topics (or sub-categories) within the selected category as part of the search terms that are sent from the banner display module to the ad search engine 230. The likelihood of the topics within the selected category can be based on the number of occurrences of the words found in the dictionary repository 250. For example, if AS/400 has the highest number of occurrences among all the other words on the page 150, the keyword analyzer 210 refines the category to "Java and AS/400".

With respect to the embodiment of FIG. 2, the keyword analyzer 210 sends the resulting refined category to the banner display module 200 which determines whether or not to display the banner ad 160 or a portion of the banner ad 160, i.e., the static portion 270, or to modify the refined category for search.

With respect to the embodiment of FIG. 3, the keyword analyzer 210 sends the resulting refined category to the web server 15 which sends the information about the refined category to the banner display module 200. In turn, the banner display module 200 determines whether or not to display the banner ad 160 or a portion of the banner ad 160, i.e., the static portion 270, or to modify the refined category for search.

As mentioned earlier, the keyword analyzer 210 can be hosted either in part on a user's computer (e.g. browser) and in part on the server 15 (FIG. 2), or on entirely on the server (15) side. The advantage of performing the analysis on the user side is to

ARC9-2000-0100-US1   24

alleviate the traffic or load on the server 15. However, in this scenario, the user's browser needs to download an executable code, such as a Java® applet and the domain specific dictionary repository 250, in order for the keyword analyzer 210 to perform the three-step analysis described above.

5

The web server 15 can be any web server that serves HTTP requests and passes them to application servers, such as the banner advertising manager 220. In turn, the banner advertising manager 220 sends the results HTTP response from the search engine 230 to the user.

10

The banner advertising manager 220 automatically constructs a query from the search terms from the keyword analyzer 210, and sends the query to the search engine 230. The query can be of any valid HTTP format, for example: "http://dw-webserver.almaden.ibm.com/cgi-bin/dWsearch.pl?UserRestriction=java," or any

15

protocol used between the banner advertising manager 220 and the search engine 230. The search engine 230 then returns the search result that contains the top 5 or more hits related to the selected categories back to the banner advertising manager 220. Using these selected categories, the banner advertising manager 220 creates an XML based file format incorporating the links and descriptions, and passes this file back to

20 the web server 15.

An exemplary XML format for the search results (hit list) based on the query "Java" in IBM® developerWorks Basic Search, illustrating five hits listed under the Community

section and enclosed in the <SR_ITEM> tags is illustrated as follows:

```
<?xml version="1.0" ?>
<SEARCH_RESULT_PAGE>
        <CATEGORIES>
            <CATEGORY_ITEM id="C1" nhits="5">Community</CATEGORY_ITEM>
        </CATEGORIES>
        <QUERY type="BasicSearch" lifetime="AdHocQuery">
                <QUERYSTRING>Java</QUERYSTRING>
        </QUERY>
        <SEARCH_RESULTS>
        <SR_ITEM category="C1">
<A HREF="http://www.software.ibm.com/ad/visage/rc/rcjav5.html"></A>
        </SR_ITEM>
        <SR_ITEM category="C1">
<A HREF="http://www.dejanews.com/dnquery.xp?search=word&amp;svcclass=dncurrent&amp;showsort=date&amp;ST=Q
S&amp;query=~g%20comp.text.xml"></A>
        </SR_ITEM>
        <SR_ITEM category="C1">
<A HREF="http://www.pageresource.com/jscript/index6.htm"></A>
        </SR_ITEM>
        <SR_ITEM category="C1">
                <A HREF="http://www.eoe.org"></A>
        </SR_ITEM>
        <SR_ITEM category="C1">
                <A HREF="http://javascript.internet.com/"></A>
        </SR_ITEM>
            </SEARCH_RESULTS>
        </SEARCH_RESULT_PAGE>
```

The search engine 230 can be any search engine capable of 26eturning the highest

ranked hits based on the input categories. For example, the IBM® developerWorks

basic search returns the top 5 hits for the domain Java with the search keyword EJB.

The search engine 230 looks into the ad repository 240 to identify the documents

related to the search terms.

It is to be understood that the specific embodiments of the invention that have been described are merely illustrative of certain application of the principle of the present invention. Numerous modifications may be made to the adaptive advertising system 10 and associated method described herein without departing from the spirit and scope of the present invention. Moreover, while the present invention is described for illustration purpose only in relation to the WWW, it should be clear that the invention is applicable as well to databases and other tables with indexed entries.